Chapter 2

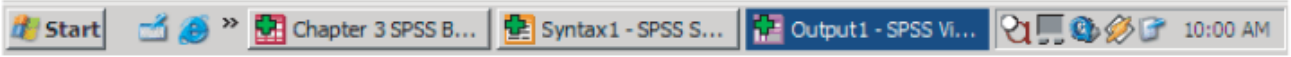
Basic Operations

Chapter 1 presented a very conceptual overview of SPSS. The present chapter will walk you through the routine procedures that you will use nearly every time you work with SPSS. Specifically you will learn how to navigate the SPSS environment, enter save and retrieve data sets; create, modify, and save text based records of procedures you preform (syntax); and generate, navigate, save and print the results of your data analyses (output).

Three Windows

For the most part SPSS can be split up into three major parts: The Data Editor (where we enter data and create new variables), The Syntax Editor (where we store & create syntax for our analyses and procedures), and The Output Navigator (where we view the results our statistical tests have generated). Each of these parts of SPSS has its own program window in SPSS. When all three are open there will be three separate buttons for SPSS on the windows task bar (See Figure 2.1a). You can navigate between the three windows by clicking on the appropriate buttons on the task bar. When working in the windows, if you close either the Output Navigator or the Syntax Window the remaining windows will not be affected. Closing the Data editor closes SPSS, and the Output Navigator and Syntax Editor will close as well.

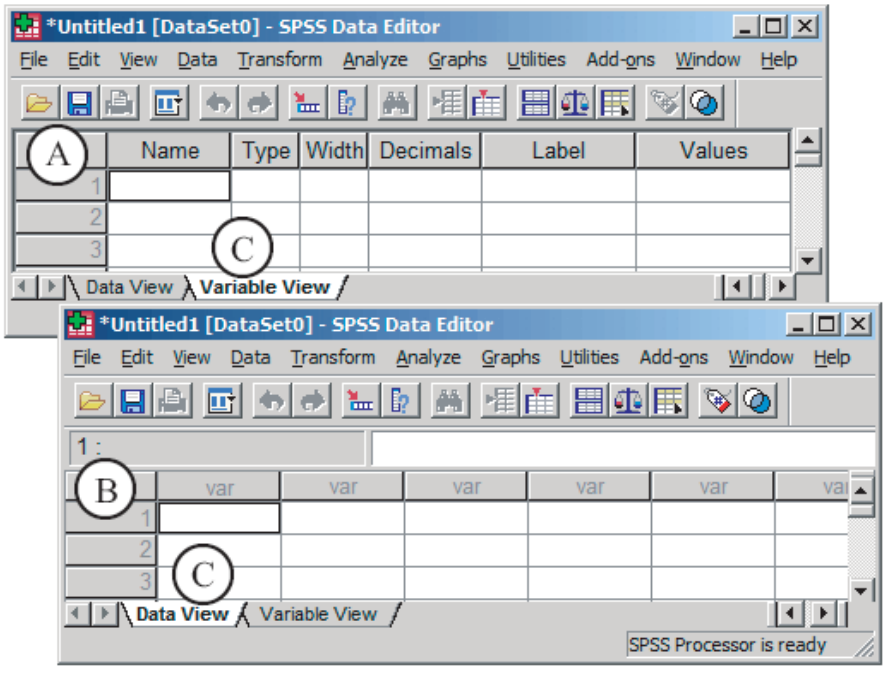
Figure 2.1a Windows Task Bar with all Three SPSS Windows Open



Data Editor

The first step in any data analysis process is to set up the data file in the Data Editor (note: you can create variables and enter data in the syntax editor, but it is beyond the scope of this text). The SPSS Data Editor is split into two parts (or views): The **Variable View** and the **Data View**. The Data View allows you to view and input data. The columns represent Variables and the rows represent observations/participants/subjects (often referred to as cases). The Variable View allows you to edit variables and add new variables to the data stet. Note that the rows represent each variable and correspond to the columns in the data view. Also, the columns in this view represent different aspects of each variable. Figure 2.1 presents the Variable View and the Data View for a blank file in the data editor. Part A represents the variable view and part B represent the data view. To toggle back and forth between the "Data View" and the "Variable View" Click on the labeled file tabs located at the bottom left hand corner of the data editor spread sheet: We have marked them with the letter "C."

Figure 2.1b SPSS: Variable View and Data View for a Blank Data File



Creating New Variables

Cartoon 2.1 is the back drop for the data set that you will work with for this chapter. Helga's statement about Hagar implies that he would not know what the well water tastes like, because he only drinks beer. Beer holds a special place in the hearts of most statisticians, as some of the basic statistical procedures and assumptions used today were developed by William Gossett, an employee of the Guinness brewery in Ireland. In the late 1800, Gossett developed the student *t*-distribution so that he could estimate the quality of a barley shipment (a population of scores) based on a random sampling from the shipment. Thus, without beer there would be no statistics. Though you may argue that without statistics there would be little need for beer. It seems likely that both statements are true.

Cartoon 3.2 HAGAR THE HORRIBLE

With respect to the data set, assume that a researcher is interested in the drinking habits of different cartoon characters. Specifically, she is interested in the types of beer consumed by cartoon characters and the amount. Type of beer would be considered a discrete (and nominal) variable that represents groups. Amount in this case is operationalized the number of beers (12 oz cans) consumed per week and will be treated as continuous (and ratio) variable. If the

distinction between discrete and continuous scores is unclear to you, go back and review the discussion in Chapter 1. Table 2.1 displays the result of the study. The first column of numbers represents the participant numbers that are used help keep track of each individual cartoon character's data. The second column presents the number of beers consumed per week by each character. The third column presents the brand of beer preferred by each character.

Table 2.1 Amount and Brands of Beer Consumed by Cartoon Characters

Participant #	*# of Beers Per Week*	*Brand*
1	1	Bongo Beer
2	2	Swiller Light
3	4	Swiller Light
4	4	Lights-Out-Lager
5	5	Lights-Out-Lager
6	6	Lights-Out-Lager
7	7	Lights-Out-Lager
8	7	Lights-Out-Lager
9	9	Budget Brew
10	10	Budget Brew
11	11	Budget Brew
12	12	Belcher's Pride
13	12	Cirrhosis Light
14	12	Cirrhosis Light
15 (Hagar's data)	15	Cirrhosis Light

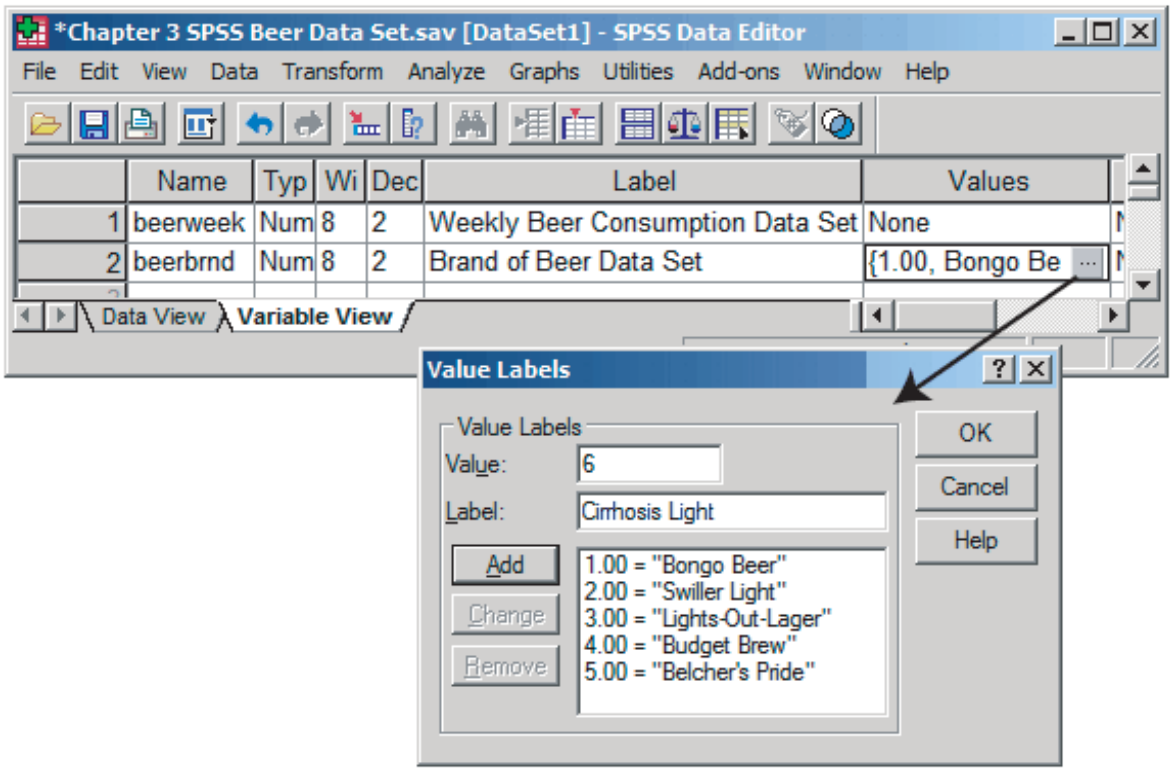
Table 2.2 Conventions of Variable Naming in SPSS 14

OK	Not OK
64 or fewer Characters	
Using Numbers	Number at Beginning of Name
Period or Decimal	Period or Decimal at End of Name
	Other Punctuation ! ? , ; : ‘ ’ “ ”
# $ _ @	% ^ & * ( ) - + = [ ] { } < > \ /

To create new variables on a blank data file, select the variable view. Note that no variables are listed. 1) Give the variable a **Variable Name** by typing the desired name in the first open row of the first column, labeled **Name**. Table 2.2 presents a summary of the variable naming do's and don'ts. SPSS 14 allows you to have 64 characters in the name, however it is a good idea to keep names as short as possible (less than 10 characters is a good rule of thumb). The Name can not start with a number, though it can have numbers in it. Nor can it have any spaces or symbols (i.e., % ^ & * + ~ ( ) { }[ ] / > <) in it, with the exception of the following: @, #, $, and _. You can use periods (or decimal points), but not at the end of the the variable name. Other forms of punctuation can not be used (! , ; : ? ‘ ’ “ ”). Figure 2.2 presents the variable view for a file where we have defined two variables. Based on Cartoon 2.1, our variables represent the

number of beers cartoon characters drink a week and the brands of beer that cartoon characters drink. In our example we have named the first variable **beerweek**. Similarly, we have named the second variable **beerbrnd**.

Figure 2.2 SPSS: Defining New Variables in the Variable View



2) Give the variable a **"Variable Label"**, by clicking the appropriate cell in the **Label** column. Variable Labels are more flexible than Variable names. You can have more than 200 characters and you can use spaces and symbols. The variable label allows you to give a more descriptive name to the variable that will make sense to you when you come back and look at your data after a long period of time has elapsed. Be as precise as possible. Also, when you give a variable a "Variable Label" the variable label will appear on the output of your analyses (we cover output later in this appendix). However, note that labels more than 40 or so characters will

be truncated on the output (it will only present the 1st 40 characters), so put the most unique and descriptive information first. Otherwise your outputs may become vary confusing. In our example we have given the **beerweek** variable the variable label "Weekly Beer Consumption Data Set," and we have given the **beerbrnd** variable the variable label "Brands of Beer Data Set."

3) If the variable represents groups (a discrete/nominal variable; e.g. sex/gender, ethnicity, group membership) then you will need to define the **Value Labels**. You do not need to do this if your data represents a continuum of scores (like 1-7 numerical rating scales: See Chapter 1 for a discussion of Levels of Measurement). By using Value Labels you can use numbers to represent groups (e.g., 1 = male, 2 = female), which makes data entry easier. Also, you will not have to try to remember what the numbers stand for, as the Value Label will be printed on the output. In our example, we will be using the numbers 1 through 6 to represent our six brands of beer.

To define Value Labels, in the **Values** column click on the row that corresponds with the variable to which you want to give value labels. When you click on this cell a small gray box will appear within the cell. Click on the box to open the **Value Labels** dialogue box, a separate window that can be closed by clicking **OK**, **Cancel**, or the small **X** in the upper right hand corner of the window. In the **Value:** field of the dialogue box type the lowest value (e.g., 0 or 1) you will be using to represent the groups. Then, in the **Value Label**: field, type the name of the category (e.g., Bongo Beer) and click the **Add** button. Next, in the **Value:** field type the next highest value (e.g., 2), then in the **Value Label:** field type the name for the next desired category (e.g., Swiller Light), and click the **Add** button. Continue this process until all the values for all

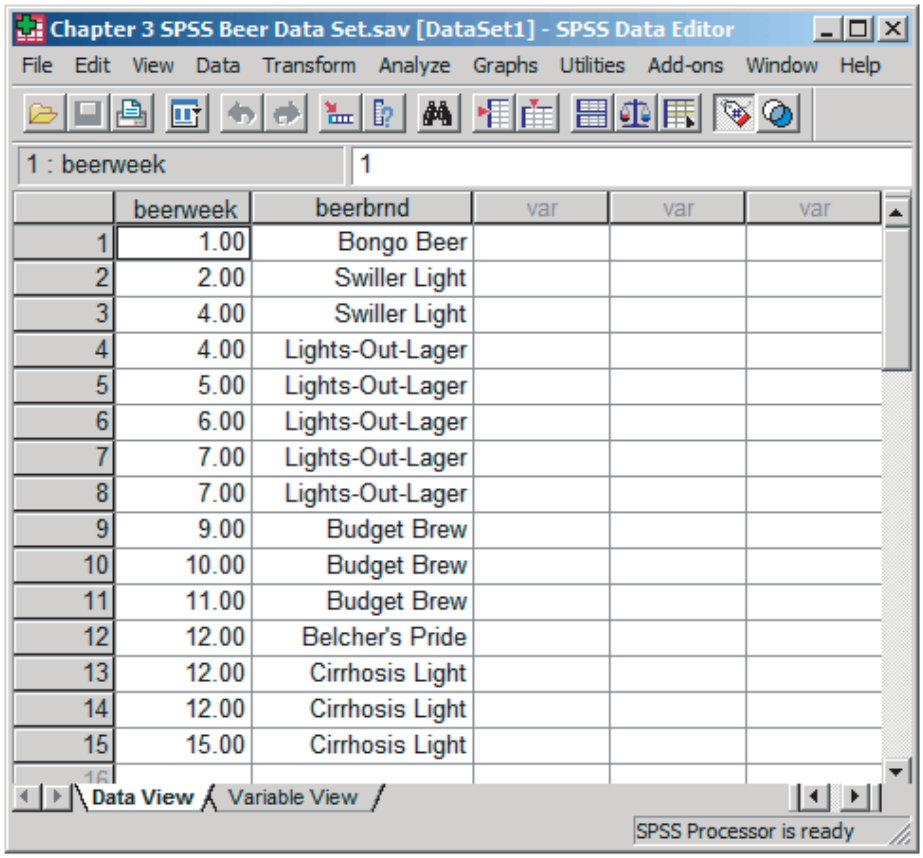
the categories have been named. When this is complete left-click **OK** and you will return the variable view of the data editor. In our example (see Figure 2.2) we have paired the following values with the following labels: 1 = Bongo Beer, 2 = Swiller Light, 3 = Lights-Out-Lager, 4 = Budget Brew, 5 = Belcher's Pride, and 6 = Cirrhosis Light.

Entering Data

Once the variables have been named and the values labeled you can begin data entry. Return to the data editor. Remember that the rows of the Data Editor represent cases/observations/subjects/participants, while the columns represent the variables you have defined. Notice that once variables are defined, the columns are labeled with the variable names. To enter data for a given case of a given variable, simply select (single left-click on) the cell you want to begin with, type the appropriate value into that cell, and then either press "enter" or one of the directional arrow keys (usually located next to the right-hand numeric key-pad). If entering data on the right-hand numeric-key-pad, be sure that the number lock has been turned on and be careful not to turn it off when reaching for the 7 key. Data can be corrected by selecting (with a left mouse click) the desired cell and typing in the new value and pressing either enter or a directional arrow key.

Figure 2.3 presents the Data View for the Beer Data Set. Here, we have entered the data for 15 cartoon characters. The first column reports the number of beers consumed per a week for each character. The second column reports the brand of beer consumed by each character. In our example, the first cartoon character drinks 1 beer a week and drinks Bongo Beer. Similarly, the 15th cartoon character (Hagar) drinks 15 beers per week and drinks Cirrhosis Light.

Figure 2.3 SPSS: Entering Data Example



	beerweek	beerbrnd	var	var	var
1	1.00	Bongo Beer			
2	2.00	Swiller Light			
3	4.00	Swiller Light			
4	4.00	Lights-Out-Lager			
5	5.00	Lights-Out-Lager			
6	6.00	Lights-Out-Lager			
7	7.00	Lights-Out-Lager			
8	7.00	Lights-Out-Lager			
9	9.00	Budget Brew			
10	10.00	Budget Brew			
11	11.00	Budget Brew			
12	12.00	Belcher's Pride			
13	12.00	Cirrhosis Light			
14	12.00	Cirrhosis Light			
15	15.00	Cirrhosis Light			

Value Labels

You should note that for variables representing groups that have had the value labels defined, the value label appears in the cell. However, when we enter the data for these variables we used the numerical values assigned to each group. For example, for subject 15, we did not type out the words "Cirrhosis Light," we simply typed 6 on the numeric keypad. It is possible to change the settings of the data editor so that it shows the numerical values instead of the value labels. This can be accomplished by selecting the **Value Labels** option from the **View** pull-down menu. When this option is checked, the value labels will appear in the data view spreadsheet. When this option is not checked, the values will appear in the data view spreadsheet.

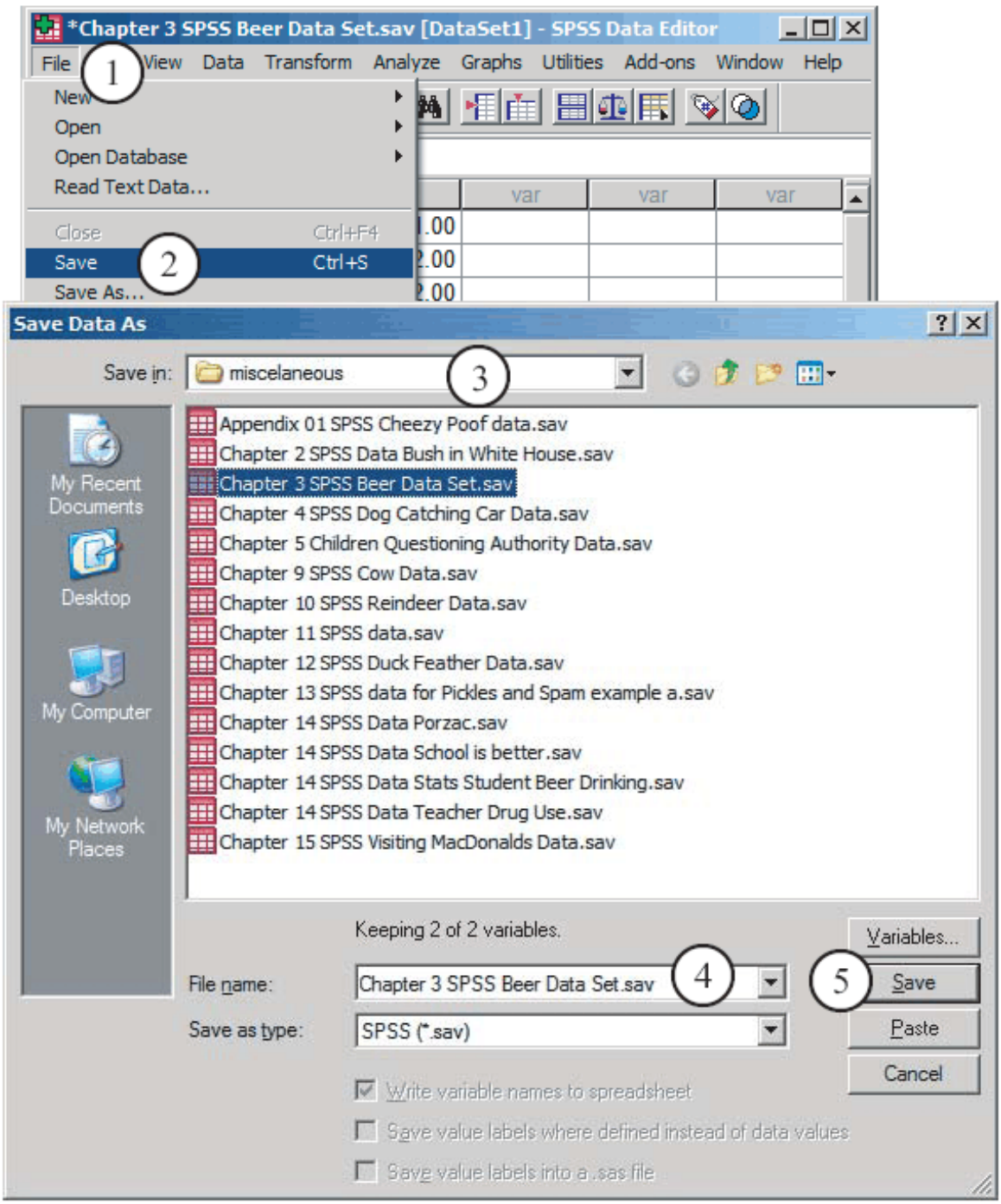
Missing Data

As a note of caution, through the course of data entry it is possible to activate data cells that you do not intend to use. For example, you may have data for 10 people, but accidently activate 11 rows. When this occurs the active cells with no data will have a decimal point in them and nothing else. When a cell is activated and there is not data entered for that cell, SPSS considers this Missing Data. Having missing data is usually not a major problem, but it can make reading the results of certain analyses a little more complicated. It is a good idea to delete those cases when you notice that you have them. This can be done by left-clicking on the row number label (located to the far left of the data editor spreadsheet) for the missing case, which will highlight the whole row. Pushing the delete button on the keyboard will deactivate all of the cells in that row. When you delete cells, be sure that you are not deleting any data that you intend to keep.

Saving Data Files

Once you have defined the variables of interest and perhaps entered some data, you should periodically save your data file. To save your data for the first time (see Figure 2.4), from the **File** (1) pull-down menu, select the **Save** or **Save as...** (2) option. In the **:Save Data As** dialogue box, select the drive [e.g., C: (the hard drive), A: (a floppy disc), etc.] and the directory (folder) that you want to save your data in by left-clicking on the boxed arow (3) on the right of the **Save in:** field and navigate through the file tree. Next, type the name you want to give the new data file in the **File name:** field (4).

Figure 2.4 SPSS: Saving Data Files Example



We should note that there are actually three types of files in SPSS: Data files, Syntax files and Output files. Each file type has a different file extension, which is the string of three letters that

appear after the period in a file name. Data files have the file extension ".sav." It is important to keep this in mind when working with different SPSS files.

Be sure to save your work often and save it in multiple places (e.g. on a hard drive or network drive and on removable media: zip disc, floppy disc, CD-R/RW, etc.) so you will not loose anything important. After you have saved your data file once, you can quickly save changes to your file by left-clicking on the save icon (the small black floppy disc) in the data editor. To save your file to different drive, a new folder, or with a different file name, you will need to use the **Save as...** option from the **File** pull-down menu.

Opening Saved Files

To open an existing data file, from the **File** pull-down menu select **Open** and select **Data...** from the side menu. In the **Open File** dialogue box, select the drive [e.g., C: (the hard drive), A: (a floppy disc), etc.] and directory (folder) that you want to retrieve your data from by left-clicking on the boxed arow on the right of the **Look in:** field and navigate through the file tree. Once you select the desired directory, double left-click on the file that you want to open or single left-click on the file and then left-click the **Open** button.

Occasionally, students can not seem to find the file they are looking for. They know it is supposed to be there, but the file does not appear in the directory as it should. You must remember that the SPSS **Open File** dialogue box will only display the files that have the file extension designated in the **Files of type:** field at the bottom of the **Open File** dialogue box. When opening data files ensure that the **SPSS (*.sav)** file type has been selected.

Syntax Files

What is Syntax?

In the "olden days" (12-15 years ago) there did not exist a Graphical User Interface (GUI) version of SPSS and all data entry and analysis was done using an SPSS syntax language. Now we have a windows based program (a GUI) and we can point & click our way through all analyses. The only problem with this is that when we want to change our analysis procedure we must step through the whole point click process again which is tiresome and potentially error ridden. However, throughout SPSS there is the option to **Paste,** which will send our point click commands to a syntax sheet that can be stored as a separate file. This way we can have all the steps in our analysis recorded and stored. Also, we can run all of our analysis from the syntax sheet, write syntax for new analyses, or alter the syntax from previous analyses.

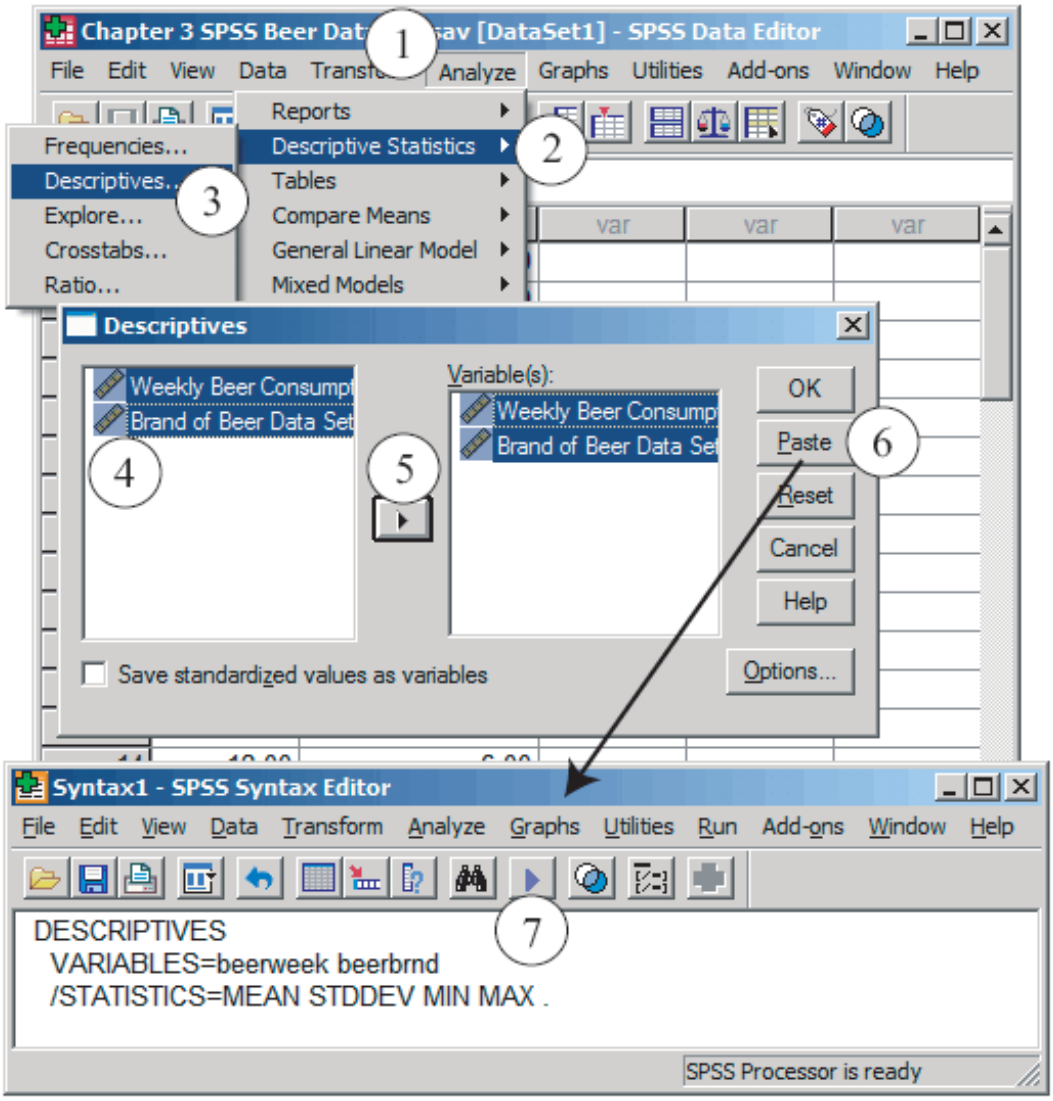
Figure 2.5 presents the steps for requesting and pasting the frequency data for our beer data. This figure also displays the resulting syntax sheet that is produced. In this example we have requested frequency analyses for the weekly beer consumption and the brands of beer variables and then have pasted the syntax by left-clicking the **Paste** button (6).

Creating, Saving and Opening a Syntax Sheet

If no syntax sheet is open when we paste our first procedure, then SPSS automatically opens an untitled sheet for us. Later, we can save the syntax as a syntax file, which has the file extension ".sps" (short for SPSS syntax file). If a syntax sheet is already open, then SPSS will paste syntax to the bottom of the currently opened syntax sheet. If multiple syntax sheets are

open, the SPSS pastes to the sheet that was opened first. To create a new syntax sheet, select **New** from the **File** pull-down menu and then select **Syntax...** from the side menu.

Figure 2.5 SPSS: Pasting Syntax Example



Once you have created a new syntax sheet you should periodically save your syntax file. To save your syntax for the first time, in the Syntax Editor, from the **File** pull-down menu select the **Save** or **Save as...** option. In the **Save As** dialogue box, select the drive [e.g., C: (the hard

drive), A: (a floppy disc), etc.] and directory (folder) that you want to save your syntax in by left-clicking on the boxed arow on the right of the **Save in:** field and navigate through the file tree. Next, type the name you want to give a the new data file in the **File name:** field (3).

To open an existing syntax file, select **Open** from the **File** pull-down menu and select **Syntax...** from the side menu. In the **Open file** dialogue box ensure that the **Files of type:** option is set to **SPSS (*.sps)**. Otherwise, the desired file may not appear as an option. Choose the desired directory from the desired drive and then select the desired file and click **Open**.

Editing Syntax

Syntax can be altered by typing, and it can be moved and removed using the copy, cut, & paste commands identical to those found in most word processors and spreadsheet programs. This becomes quite useful when you are running the same type of analysis many times with only slight alterations. For example, you may want to run the same statistical test multiple times, but use a different dependent variable each time. You could point-click your way through each desired analysis, or you could point-click your way through one analysis, paste it, copy the syntax, paste the copied syntax (using the paste command from the Edit pull-down menu), and then change the variable name in the syntax by hand. It may take you some time to become comfortable working with syntax files, but once you do, it will really pay off.

Running Analyses From Syntax

To run an analysis or multiple analyses from the syntax box, highlight the syntax for the desired analyses and then press the play button on the tool bar. The **play** button (**7**) resembles the

play button on a VCR or CD player (arrow pointing to the right). Also, analyses can be run from the **Run** pull-down menu, where there are four options: **All** (runs all analyses on the currently active syntax sheet), **Selection** (runs highlighted area), **Current** (runs the analysis your cursor is currently positioned within), and **To End** (runs analyses on syntax sheet that fall below the point where the cursor is currently positioned). Finally, you can run selected syntax by highlighting the desired analyses and right-clicking on the selected syntax. From the list of options that appear in the pop-up menu, select **Run Current**.

Output Files

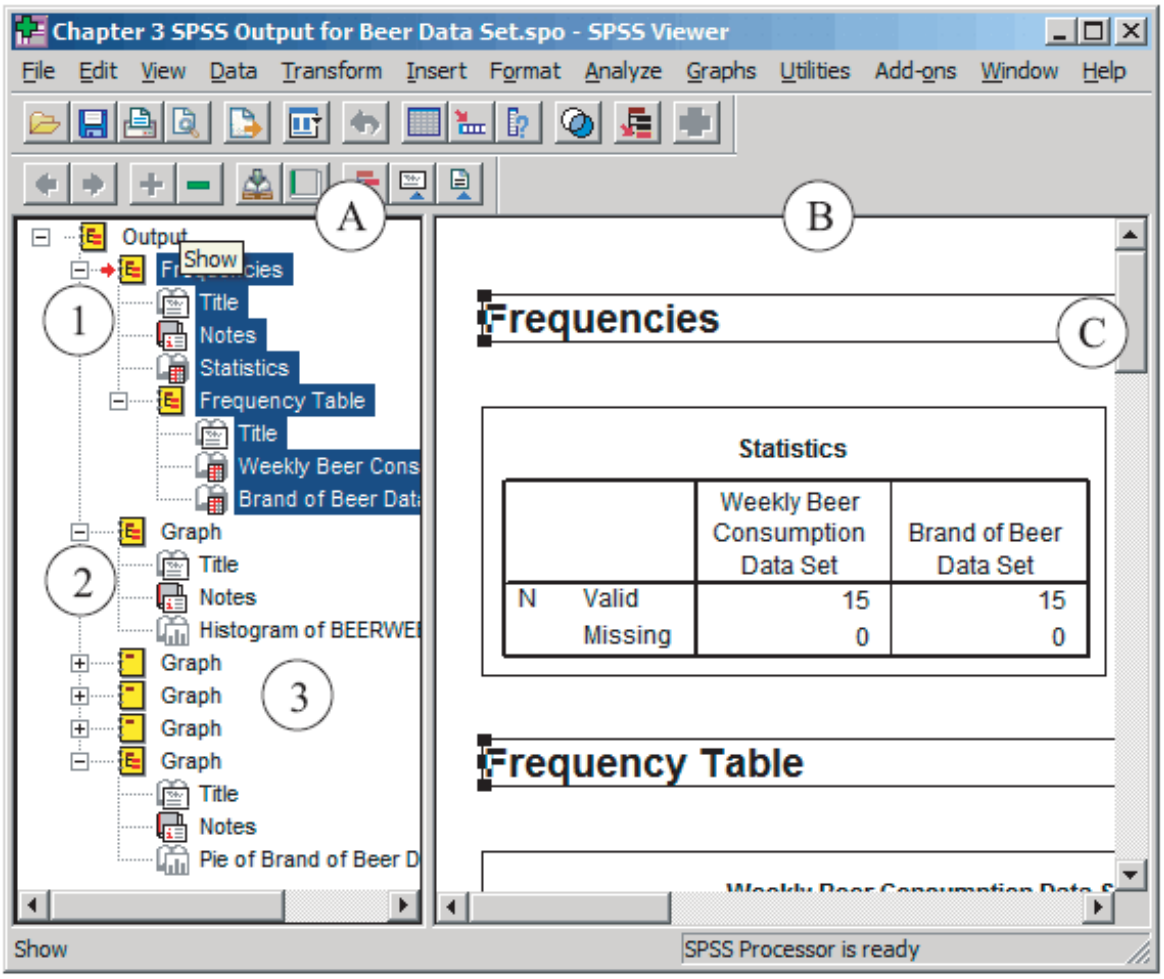
Anytime you run an analysis (e.g., Desriptives, Frequencies, Chi-Square, etc.) the results will be presented in a separate window titled Output Viewer. Like syntax, this is a separate file that will need to be saved. The extension for these types of files are ".spo" (short for SPSS output). Note that all analyses can be run from the output window, but some options like the Data or Transform can not be requested (This is not true of the newer versions of SPSS where they have made the options consistent in the data editor, syntax editor, and output navigator).

Navigating the Output

Figure 2.6 presents an output file where six different analyses have been requested. The Output Navigator is split into two parts: The Outline View (labeled A in Figure 2.6) and the Output Display (labeled B in Figure 2.6). The outline view allows us to quickly move from one analysis output to another by clicking on the part of the analysis you wish to view. For example,

in Figure 2.6, the Frequency Analysis (1) is selected and the first part of the frequency analysis (labeled Statistics) is presented in the Output Display field. Similarly, you can move to the first graph you requested by clicking on the Graph analysis heading (2) of the Outline.

Figure 2.6 SPSS: Navigating the Output Viewer



The Outline View also allows you to collapse and expand parts of analyses or a whole analysis by double-left-clicking on the desired analysis heading. For example, in Figure 2.6 three of the Graph analyses (3) have been collapsed. When analyses or parts of analyses are collapsed, they will not appear in the Output Display field, nor will they appear on printouts of the output

(unless you change the print settings). Collapsed parts of the output can be re-expanded by double-left-clicking on the collapsed analysis heading in the Outline View. Analyses can also be deleted from the output file by selecting the desired analysis heading and then using the delete button on the keyboard. You can also copy and paste analyses to different parts of the Outline in order to change the order in which they appear in the Output Display and in any printouts produced.

The Output Display can also be used to navigate through the output that you generate. By Clicking and dragging the Scroll Bar (C) to the right of the Output Display, you can move up and down through the output. Also, by left-clicking a part of the output in the Output Display, it becomes active which highlights that part of the analysis in the Outline View. This is a quick way to make sure that you are really looking at the output that you want. It is especially useful when you have conducted the several similar analyses with similar variables.

Printing Output

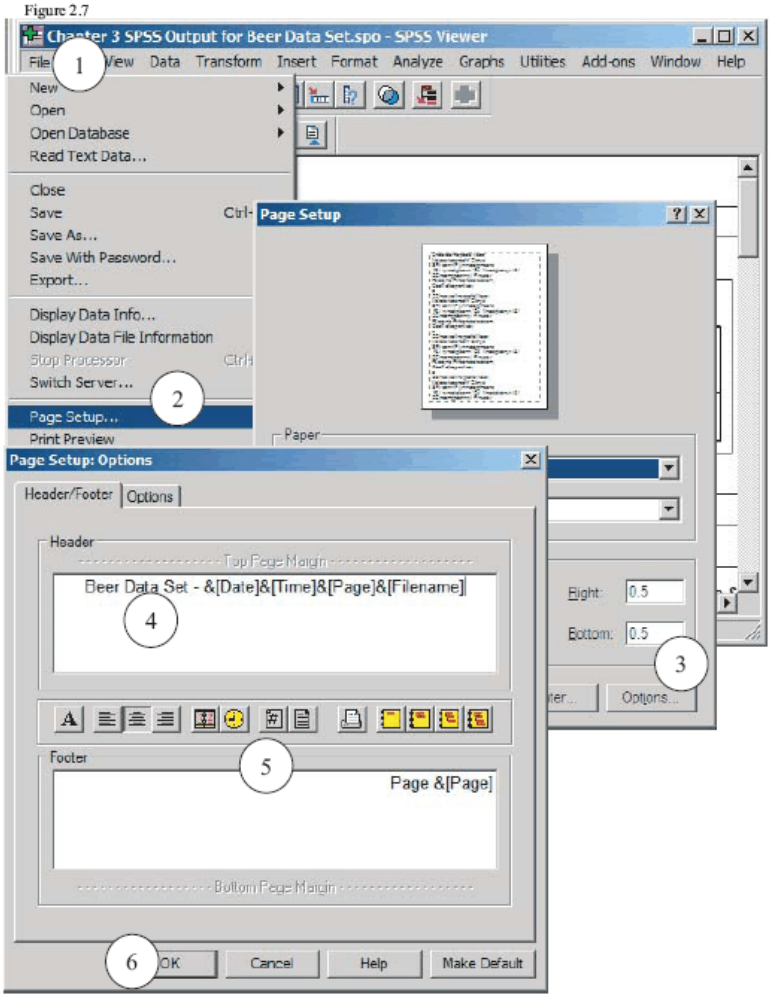
To print the output, first ensure that you have all the desired analyses expanded (and all undesired analyses collapsed) in the Outline View. You can use either printer icon on the Output Navigator toolbar, or you can select the **Print...** option from the **File** pull-down menu. In the **Print** window ensure that the **All visible output** option has been selected. This option prevents analyses that you have collapsed (and therefore hidden) from printin. Click **Print** to finish.

Adding a Header to your printouts

Having hard copies of your output is quite useful. However, you will often print multiple

versions of your analysis output. Especially if you find mistakes or add more cases/people to your data set. It is helpful to be able to label your output with identifying information that keeps you from mixing output pages of other analyses together. You can label your output by using a Page Header. In a page header, you can include the time and date the analysis was printed, the filename of the output file, and any custom information you would like to add. Figure 2.7 illustrates the steps involved.

Figure 2.7



Adding a Page Header

Steps (See Figure 2.7): In the Output Navigator, from the **File** (1) pull down menu, select **Page Setup...** (2). In the **Page Setup** dialogue box, left click **Options...** (3). In the **Page Setup: Options** dialogue box, add the Date, Time, Page Number, and Filename by clicking the icons (4) below the Header field. You can also choose whether the header will be centered, flush right (right justified) or flush Left (left justified). Custom information can be included by typing in the header field. In this example the custom information (Beer Data Set -), the date, the time, page number, and filename have been

requested. To finish, left click **OK** (6).

The header you add will now appear at the top of every page of the printed output. However, it will not appear anywhere on the screen, when you are working with the output navigator.

Saving and Opening Output

Once you have generated output and ensured that your analyses were run correctly, you should save the output file. To save your output for the first time, in the Output Editor, from the **File** pull-down menu, select the **Save** or **Save as...** option. In the **Save As** dialogue box, select the drive [e.g., C: (the hard drive), A: (a floppy disc), etc.] and directory (folder) that you want to save your output in by left-clicking on the boxed arow on the right of the **Save in:** field and navigate through the file tree. Next, type the name you want to give a the new output file in the **File name:** field.

To open an existing Output file, select **Open** from the **File** pull-down menu and select **Output...** from the side menu. In the **Open File** dialogue box ensure that the **Files of type:** option is set to **SPSS (*.spo)**. Otherwise, the desired file may not appear as an option. Choose the desired directory from the desired drive and then select the file you want and click **Open**.

Chapter Summary

This chapter introduced many of the most common operations used in SPSS: creating and retrieving data files, working with syntax files, and working with SPSS output. As you progress to more advanced chapters you may wish to periodically review the procedures outlined here.

Also, you may discover that there are usually several different ways to accomplish the same objective (some you may like, some you may not). There is no right or wrong way to do things, as long as the procedure does what you, and those you work with, need done. There is space here at the end of the chapter to make notes about the procedures you use most often and the ones you discover. It will be helpful for you to keep track of what works for you, what does not, and why.